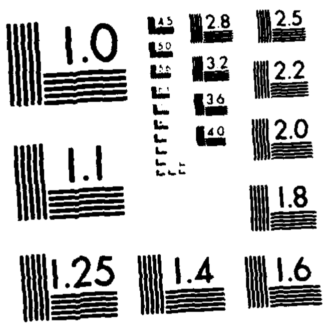
AD-A136 487 A COMPARISON OF RANDOM BALANCE AND TWO-STAGE GROUP SCREENING DESIGNS PART 1(U) DESMATICS INC STATE COLLEGE PA C A MAURO DEC 83 TR-113-12 N00014-79-C-0650 1/1

UNCLASSIFIED F/G 12/1 NL

END
DATE
FILMED
84
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

A COMPARISON OF RANDOM BALANCE AND TWO-STAGE GROUP SCREENING DESIGNS - PART I

by

Carl A. Mauro

12

DESMATICS, INC.

P. O. Box 618
State College, PA. 16801
Phone: (814) 238-9621

Applied Research in Statistics - Mathematics - Operations Research

A COMPARISON OF RANDOM BALANCE AND TWO-STAGE GROUP SCREENING DESIGNS - PART I

by

Carl A. Mauro

TECHNICAL REPORT NO. 113-12

December 1983

This study was supported by the Office of Naval Research under Contract No. N00014-79-C-0650, Task No. NR 042-467

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	1
2.	MODEL ASSUMPTIONS	2
3.	THE RB AND GS STRATEGIES	3
	3.1. Random Balance	4
	3.2. Two-Stage Group Screening	4
	3.3. Performance Measures	5
4.	RESULTS OF COMPARISON STUDY	6
5.	DISCUSSION	8
6.	PRACTICAL CONSIDERATIONS	13
7.	CONCLUSIONS AND SUMMARY	16
8.	APPENDIX	19
9.	REFERENCES	21

1. INTRODUCTION

Computer simulation is often used to study real-world systems that are too complex to be modeled and analyzed entirely by mathematical methods. Unfortunately, simulation models used to study large, complex systems tend to be extremely large and complex themselves, and corresponding computer codes (programs) to execute these models are generally very long running. As a consequence, users of large-scale simulations are often overwhelmed by the vast number of factors (i.e., input variables) contained in the model and are confused about how to make an effective analysis of the system model without having to perform an excessive number of costly and time-consuming simulation runs. Methods of shortcutting these cost and time elements are essential if any fruitful simulation experimentation is to take place. In such situations, the use of factor screening methods can substantially reduce the total number of computer runs required to study the system model.

Factor screening methods (see, for example, [1], [7], and [9]) are statistical methods that attempt to identify, economically and efficiently, a set of "most important" factors. Once the more important factors have been identified, subsequent simulation experimentation can be focused on these critical factors, thereby eliminating experimentation with relatively negligible factors which can needlessly consume resources. Although factor screening methods are applicable to experimentation in general, computer simulation offers an especially fertile area of application for these techniques for at least two reasons: (a) the large

number of factors generally built into complex simulation models, and (b) the scarcity of computer runs that often handicaps planned simulation experimentation.

Under an Office of Naval Research contract, Desmatics, Inc. has conducted extensive research in this area. As part of its research effort, Desmatics has selected two primary screening strategies for intensive study. These two strategies are Random Balance (RB) and Two-Stage Group Screening (GS). In earlier technical reports the respective performance characteristics of these two strategies were evaluated. The present report is the first of two technical reports to compare directly the performance of the RB and GS strategies.

2. MODEL ASSUMPTIONS

Suppose that K factors, each at two levels (±1), are to be screened for their effect on the simulation response (i.e., output variable). For detecting the factors having major effects it is generally reasonable to assume the first-order model

$$y_i = \beta_0 + \sum_{j=1}^{K} \beta_j x_{ij} + \varepsilon_i \tag{2.1}$$

where y_i is the value of the response in the $i^{\underline{th}}$ simulation run, x_{ij} is the level (±1) of the $j^{\underline{th}}$ factor in the $i^{\underline{th}}$ simulation run, β_j is the (linear) effect of the $j^{\underline{th}}$ factor, and the ε_i are i.i.d. $N(0,\sigma^2)$ random disturbances, $\sigma^2 > 0$ unknown. Ordinarily we would use model (2.1) over a relatively small region of the factor space.

In this report we make the following additional simplifying assumptions:

(i) $k \geq 1$ (k unknown) of the K factors are active (i.e., have a nonzero effect) and the remaining (K-k) factors are inactive, and

(ii) all active factors have the same absolute effect, $\Delta > 0$, that is,

$$|\beta_j| = \begin{cases} \Delta, \text{ if the } j^{\underline{th}} \text{ factor is active} \\ 0, \text{ if the } j^{\underline{th}} \text{ factor is inactive.} \end{cases}$$

We let $\underline{\beta}(i)$ for $i = 0,1,\ldots,k$ denote the case in which (of the k active effects) i active effects are equal to $-\Delta$ and the remaining (k-i) active effects are equal to $+\Delta$. The (K-k) inactive effects are, by definition, equal to zero. We note that the $\underline{\beta}(0)$ case, or $\underline{\beta}(k)$ case, corresponds to the situation in which all k active effects are in the same direction. In practice, of course, the direction will be known for some suspected effects and unknown for others. Lastly, we define the signal-to-noise ratio as the ratio of Δ to the error standard deviation σ, Δ/σ.

3. THE RB AND GS STRATEGIES

In this section we review briefly the RB and GS strategies. These strategies are discussed more fully in [2] and [6]. In addition, we define three basic measures of performance which we will use in Section 4 to compare screening performance.

3.1. Random Balance

In a two-level (±1) RB design for studying K factors, each column of the design matrix consists of N/2 +1's and N/2 -1's where N (an even number) denotes the total number of runs to be made. In each design column the +1's and -1's are allocated randomly, making all possible combinations of N/2 +1's and N/2 -1's (there are $C^N_{N/2}$ in all) equally likely, with each column receiving an independent randomization. To analyze RB designs we apply a standard F-test separately to each factor, ignoring all other factors. Furthermore, we conduct each F-test at the same level of significance, say α_r. Our RB strategy, therefore, is completely specified by N and α_r. Accordingly, we denote such a screening strategy by RB(N, α_r).

3.2. Two-Stage Group Screening

In this strategy we partition the K factors randomly into G groups of size g; if K is not a multiple of g, we assume that the group sizes are taken as "evenly" as possible. Then, by assigning the same level (+1 or -1) to all component factors within each group, we test the group factors as if they were single factors. All factors in groups found to have a significant effect are subsequently studied in a second-stage experiment.

In the first and second stage experiments we use the multifactorial designs of Plackett and Burman [8]. These designs are specially constructed two-level orthogonal desings for studying up to (4m-1) factors

in 4m runs. The number of runs required by the smallest Plackett-Burman (PB) design to study s factors is given mathematically by

$$B(s) = s+4 - s(\bmod 4).$$

PB designs can be analyzed by the usual analysis of variance procedures for factorial experiments. Note, however, that when the number of factors is one less than a multiple of four, no degrees of freedom are left to estimate experimental error (σ). It is advisable, therefore, for the study of s factors (or group-factors) to employ the PB design in B(s+1) runs. This would result in a minimum of one and a maximum of four error degrees of freedom.

If we let α_1 and α_2 denote the levels of the significance tests performed at the end of the first and second stages, respectively, our GS strategy is completely specified by g, α_1, α_2. We denote such a strategy by GS(g, α_1, α_2).

3.3. Performance Measures

With regard to model (2.1) and the additional simplifying assumptions we have made, we can define three basic measures of screening effectiveness. These are:

> Power. We denote by A the number of active factors that are detected correctly, and we define $E_A = 100E(A)/k$ as a percentage measure of the power (or sensitivity) to detect the active factors.

Type I Error. We denote by U the number of inactive factors that are declared active, and we define $E_U = 100E(U)/(K-k)$ as a percentage measure of Type I error.

Relative Testing Cost. We denote by R the total number of runs required by an RB or GS strategy. We define $E_R = 100E(R)/B(K+1)$ as a percentage measure of expected relative testing cost, where B(K+1) denotes the number of runs required by a PB design for (K+1) factors.

In references [2], [3], and [6], formulas are given with which to calculate E_A, E_U, and E_R for any GS(g, α_1, α_2) or RB(N, α_r) strategy. We apply these results, as needed, in the following section in making our comparative performance study.

4. RESULTS OF COMPARISON STUDY

In our investigation we examined the following eight combinations of $(K, k, \Delta/\sigma)$:

$$K = 60 \text{ and } 240,$$

$$k = (1/15)K \text{ and } (4/15)K,$$

$$\text{and} \quad \Delta/\sigma = 2 \text{ and } 4.$$

Further, for each of these eight cases we considered the following eight combinations of Type I Error and expected number of runs:

$$E_U = 10\% \text{ and } 20\%$$

$$\text{and } E(R) = \begin{cases} 12,\ 26,\ 38,\ 52 \text{ runs} & \text{for } K=60 \\ 46,\ 100,\ 144,\ 198 \text{ runs} & \text{for } K=240. \end{cases}$$

We chose these particular run numbers to correspond as closely as possible to expected relative testing costs (E_R) of 20%, 40%, 60% and 80%. We could not specify run numbers that gave exact correspondence between E_R and these levels. The reason for this is that in an RB(N, α_r) strategy the total number of runs required, R, is constant and is precisely N, which must of course be kept even. Consequently, E(R) had to be restricted to even numbered runs only. A quick calculation will show that for K=60 factors, we consider E_R = 18.75%, 40.63%, 59.38%, and 81.25%; for K=240 factors, we consider E_R = 18.85%, 40.98%, 59.02%, and 81.15%.

In all, then, we considered 64 treatment combinations of K, k, Δ/σ, E_U, and E_R. For each treatment combination, we determined that RB(N, α_r) and GS(g, α_1, α_2) strategies that maximize power E_A, and thus are "optimal" in this sense. For GS strategies we did this separately for the $\underline{\beta}(0)$ and $\underline{\beta}([k/2])$ cases. The $\underline{\beta}(0)$ case represents the "best" case situation for group screening since no cancellation of active effects can occur within groups. The $\underline{\beta}([k/2])$ case, on the other hand, represents the "worst" case situation for group screening since in this case the chance of group-factor cancellation of active effects is greatest (among all $\underline{\beta}(i)$ cases). In contrast, the sensitivity of an RB(N, α_r) strategy is the same for all $\underline{\beta}(i)$ cases, i = 0,1,...,k.

The optimal RB(N, α_r) strategy for a given (K, k, Δ/σ) condition can be quite readily determined. Mauro and Smith [6] have shown that the Type I error of an RB(N, α_r) strategy is very closely approximated by α_r. Furthermore, the power of an RB(N, α_r) strategy increases as either N or α_r increases. It follows that for a given (K, k, Δ/σ, E_U, E_R)

condition the optimal RB(N, α_r) strategy is simply the RB(E(R), E_U) strategy.

The optimal GS(g, α_1, α_2) strategy is much more difficult to determine. Mauro [2] and Mauro and Burns [4], however, have developed a computer-aided search routine that can be used to determine the optimal GS(g, α_1, α_2) strategy under the same model assumptions we have made in this paper. Moreover, the algorithm treats both the $\underline{\beta}(0)$ and $\underline{\beta}([k/2])$ cases. Accordingly, this search routine was applied to determine the optimal GS(g, α_1, α_2) strategy in each of the 64 experimental conditions.

The corresponding powers of the optimal RB(N, α_r) and GS(g, α_1, α_2) strategies are presented for easy comparison in Tables 1 and 2. Table 1 summarizes the results for K=60 factors and Table 2 does so for K=240 factors. The values of g, α_1, α_2 that define the optimal GS(g, α_1, α_2) strategies are not given in these tables but are listed in the appendix. For notational purposes and convenience of presentation, we define GS.i for i = 0,1,...,k to be the optimal GS(g, α_1, α_2) strategy in the $\underline{\beta}(i)$ case. Although we only consider i=0 and i=[k/2], it is clear because of symmetry considerations that the GS.i strategy is equivalent to and has the same power as the GS.(k-i) strategy.

5. DISCUSSION

As noted previously, $\underline{\beta}(0)$ and $\underline{\beta}([k/2])$ are the "best" and "worst" case situations, respectively, for group screening. Consequently, the power corresponding to the GS.0 strategy is always greater than that

Type I Error	Strategy	Expected Number of Runs 12	26	38	52
	RB	68.8[1]	92.6	98.0	99.6
20%	GS.0	9.2	54.1	97.7	100.0
	GS.2	7.5	46.0	90.1	96.9
	RB	53.0	85.4	95.2	98.8
10%	GS.0	9.2	54.1	97.7	100.0
	GS.2	7.5	46.0	90.1	96.9

(a) K=60, k=4, $\Delta/\sigma=2$

Type I Error	Strategy	Expected Number of Runs 12	26	38	52
	RB	33.8[1]	49.6	60.4	70.4
20%	GS.0	7.9	30.9	56.6	82.1
	GS.8	7.5	26.5	47.0	69.7
	RB	20.5	34.6	45.4	56.4
10%	GS.0	7.9	30.9	56.6	82.1
	GS.8	7.5	26.5	47.0	69.7

(b) K=60, k=16, $\Delta/\sigma=2$

Type I Error	Strategy	Expected Number of Runs 12	26	38	52
	RB	70.5[1]	93.7	98.4	99.7
20%	GS.0	9.3	55.4	100.0	100.0
	GS.2	7.5	47.5	90.6	97.2
	RB	54.9	87.1	96.1	99.1
10%	GS.0	9.3	55.4	100.0	100.0
	GS.2	7.5	47.5	90.6	97.2

(c) K=60, k=4, $\Delta/\sigma=4$

Type I Error	Strategy	Expected Number of Runs 12	26	38	52
	RB	34.0[1]	49.9	60.8	70.8
20%	GS.0	7.9	31.0	56.7	82.2
	GS.8	7.5	26.5	47.0	69.7
	RB	20.6	34.9	45.8	56.8
10%	GS.0	7.9	31.0	56.7	82.2
	GS.8	7.5	26.5	47.0	69.7

(d) K=60, k=16, $\Delta/\sigma=4$

[1] Powers are expressed as percentages

Table 1. Power Comparisons of Optimal RB and GS Strategies for K=60 Factors.

Type I Error	Strategy	Expected Number of Runs 46	100	144	198
	RB	66.4[1]	89.5	96.2	98.9
20%	GS.0	22.8	69.0	100.0	100.0
	GS.8	17.6	61.9	94.1	97.3
	RB	51.9	81.2	92.0	97.4
10%	GS.0	22.8	69.0	100.0	100.0
	GS.8	17.6	61.9	94.1	97.3

(a) K=240, k=16, $\Delta/\sigma=2$

Type I Error	Strategy	Expected Number of Runs 46	100	144	198
	RB	34.5[1]	49.2	58.9	68.5
20%	GS.0	14.9	41.7	63.8	85.4
	GS.32	13.7	37.0	55.4	78.2
	RB	21.4	34.6	44.2	54.6
10%	GS.0	14.9	41.7	63.8	85.4
	GS.32	13.7	37.0	55.4	78.2

(b) K=240, k=64, $\Delta/\sigma=2$

Type I Error	Strategy	Expected Number of Runs 46	100	144	198
	RB	66.8[1]	89.8	96.3	99.0
20%	GS.0	22.9	69.1	100.0	100.0
	GS.8	17.6	62.0	94.1	97.3
	RB	52.3	81.6	92.2	97.5
10%	GS.0	22.9	69.1	100.0	100.0
	GS.8	17.6	62.0	94.1	97.3

(c) K=240, k=16, $\Delta/\sigma=4$

Type I Error	Strategy	Expected Number of Runs 46	100	144	198
	RB	34.5[1]	49.2	59.0	68.6
20%	GS.0	15.0	41.7	63.8	85.5
	GS.32	13.7	37.0	55.4	78.2
	RB	21.5	34.6	44.3	54.7
10%	GS.0	15.0	41.7	63.8	85.5
	GS.32	13.7	37.0	55.4	78.2

(d) K=240, k=64, $\Delta/\sigma=4$

[1]Powers are expressed as percentages.

Table 2. Power Comparisons of Optimal RB and GS Strategies for K=240 Factors.

corresponding to the GS.[k/2] strategy. As seen from Tables 1 and 2, this difference in power becomes greater as expected relative testing cost increases.

Further inspection of Tables 1 and 2 reveals that Δ/σ, over the range we considered, had little effect on the powers of the optimal RB strategies and virtually no effect on the powers of the optimal GS strategies. Effectively, therefore, we can ignore Tables 1c, 1d, 2c, and 2d and restrict attention to Tables 1a, 1b, 2a, and 2b (or vice-versa). We suspect, though, that had we considered a signal-to-noise ratio somewhat smaller than two (say $\Delta/\sigma = 1$), there would have been some loss in power compared with $\Delta/\sigma = 2$ and 4. However, signal-to-noise ratios less than two are probably not of practical interest in screening situations.

From a comparison of Tables 1a with 2a and 1b with 2b, it is readily seen that the powers of the optimal RB strategies depend on K and k basically only through $p = k/K$, the proportion of active factors to the total number of factors. This simple relationship apparently does not hold for GS strategies.

In this study p ranges from 6.7% (=1/15) to 26.7% (=4/15). We see from the tables that the optimal RB and GS strategies have greater power when p = 6.7%. This observation is in accordance with the notion that factor screening is more effective for a given K in the presence of fewer active factors (i.e., for smaller p). It can also be seen from the tables that the drop in power as p increases from 6.7% to 26.7% is more extreme for the optimal RB strategies than for the optimal GS strategies.

Continuing, it is clear that both Type I error and relative testing cost have a strong influence on the powers of the optimal RB strategies.

For the optimal GS strategies, relative testing cost similarly has a major effect on power. Type I error, however, has little effect on the powers corresponding to optimal GS strategies. This result was somewhat unexpected and we investigated this phenomenon further for a few selected conditions. Surprisingly for these cases we found relatively little loss in power for the optimal GS strategies with Type I error rates as low as 1%. On the other hand, small Type I error rates generally have a debilitating effect on the use of RB strategies.

For the remainder of this section we shall attempt to discuss the relative merits of each screening strategy. This discussion should provide some guidance and insight into the use and selection of these two techniques for factor screening.

A primary question of interest is for what combinations of Type I error and expected relative testing cost should one consider the use of an RB strategy rather than a GS strategy. Over the range of Type I error considered, the optimal RB strategy is "better" (i.e., has greater power) for low expected relative testing cost, and the optimal GS strategy is better for high expected relative testing cost. In the $\underline{\beta}(0)$ case, the crossover for $E_U = 10\%$ is about $E_R = 45\%$ and for $E_U = 20\%$ the crossover is about $E_R = 60\%$. The crossover, however, varies widely with K and k. In general, the crossover decreases (increases) as either K or k decreases (increases). In the $\underline{\beta}([k/2])$ case the crossover shifts upward about 15%.

Preliminary extrapolation studies have indicated that the E_R crossover, where the optimal GS strategy becomes better than the optimal RB strategy, is smaller (larger) for smaller (larger) levels of Type I

error. This would therefore suggest that the optimal GS strategy has an advantage over the optimal RB strategy at low Type I error rates but begins to lose this advantage as one considers screening at higher Type I error rates. Of course, in a particular screening application, it is up to the analyst to make the appropriate compromises between Type I error, relative testing cost, and power.

There are two very important practical considerations that should be noted. The first of these is that the total number of screening runs required by an RB(N, α_r) strategy is fixed prior to experimentation. In an GS(g, α_1, α_2) strategy, the total number of runs required is random. The RB strategy, therefore, offers greater control over the number of screening runs that will be expended.

The second and perhaps most important consideration is that for a given expected relative testing cost, determination of the optimal GS(g, α_1, α_2) strategy requires prior knowledge of k, Δ/σ, and the number of active effects in the positive direction. On the other hand, determination of the optimal RB(N, α_r) strategy does not require this, or any other, prior knowledge. Consequently, any advantages of the optimal GS strategies (as indicated in the tables and discussed so far) may be offset by losses in power due to imprecise prior knowledge. We examine this potential hazard in more detail in the following section.

6. PRACTICAL CONSIDERATIONS

A desirable feature of any factor screening procedure is the ability to control Type I error and expected relative testing cost. As indicated

previously, this is always possible with an RB(N, α_r) strategy. With a GS(g, α_1, α_2) strategy, however, this control is possible only with prior knowledge of k, Δ/σ, and $\underline{\beta}(i)$. An important question, therefore, is to what extent does imprecision in this prior knowledge affect the performance of a GS(g, α_1, α_2) strategy. In this section we attempt to answer this question through the use of two case studies. These examples will serve to illustrate the practical difficulties associated with the application of the GS strategy.

In the first case study, we assume that there are K=60 factors to be screened for their effect on the response. In addition, suppose that we wish to control our Type I error at 10% and expected relative testing cost at 59.4% (equivalently, E(R) = 38 runs). Further suppose that our prior knowledge tells us to expect that k=16 factors are active, Δ/σ = 2, and all active effects are in the same direction. We see from the appendix that the optimal GS strategy for this situation is the GS(7, 0.00325, 0.29673) strategy. Suppose for the moment, however, that our prior knowledge is not entirely accurate. In Table 3 we give the performance of the GS(7, 0.00325, 0.29673) strategy for all combinations of k, Δ/σ, and $\underline{\beta}(i)$ for k=8,12,16, Δ/σ=2,4 and $\underline{\beta}(i)=\underline{\beta}(0)$, $\underline{\beta}([k/2])$.

In the second case study we assume that K=240 factors are to be screened. Once again suppose that we wish to control Type I error at 10% and expected relative testing cost at 59.0% (equivalently, E(R) = 144 runs), and suppose our prior knowledge tells us to expect that k=16, Δ/σ=4, and $\underline{\beta}(i)=\underline{\beta}(0)$. From the appendix, the optimal GS strategy for this situation is the GS(3, 0.05907, 0.55223) strategy.

$\underline{K}$	$\underline{k}$	$\underline{\Delta/\sigma}$	$\underline{\beta(i)}$	$\underline{E}_U$	$\underline{E}_R$	$\underline{E}_A$
60	8	2	$\beta(0)$	4.4	38.9	35.6
60	8	2	$\beta(4)$	2.9	32.7	18.0
60	8	4	$\beta(0)$	10.4	59.4	71.5
60	8	4	$\beta(4)$	7.8	49.3	44.6
60	12	2	$\beta(0)$	7.1	49.0	46.6
60	12	2	$\beta(6)$	4.0	36.6	20.8
60	12	4	$\beta(0)$	14.9	74.6	80.8
60	12	4	$\beta(6)$	10.1	56.7	47.1
60	16	2	$\beta(0)$	10.0	59.4	56.6
60	16	2	$\beta(8)$	5.0	40.1	23.4
60	16	4	$\beta(0)$	18.7	87.0	87.5
60	16	4	$\beta(8)$	11.9	62.2	49.6

Table 3. Performance Results for GS(7, 0.00325, 0.29673) Strategy. All Results Are Expressed as Percentages.

In Table 4 we give the performance of this particular GS strategy for all combinations of k, Δ/σ, and $\underline{\beta}(i)$ for k=16,24,32, Δ/σ=2,4, and $\underline{\beta}(i)=\underline{\beta}(0),\underline{\beta}([k/2])$.

As can be seen from Tables 3 and 4, Type I error and expected relative testing cost can deviate greatly from their intended values, although they always move in the same direction. From a practical standpoint, these results indicate that the use of imprecise prior knowledge can have rather undesirable consequences. Underestimating the number of active factors results in greater type I error and greater expected relative testing cost than desired. Overestimating the number of active factors has the reverse effect. Certainly, this is a major practical drawback to group screening as a technique for factor screening.

7. CONCLUSIONS AND SUMMARY

In this paper we attempt to compare the efficacy and relative merits of a two-stage group screening (GS) strategy versus a random balance (RB) screening strategy. We assume a screening model in which the active (i.e., nonzero) effects are additive and have the same absolute magnitude. Accordingly, this model is most appropriate when it is expected that a relatively small number of factors (i.e., inputs) have a major effect on the response (i.e., output) and the remaining factors have a negligible effect. In such situations, the objectives of a screening strategy are to detect as many of the "important" factors as possible, to declare important as few "unimportant" factors as possible, and to perform as few computer runs as possible.

$\underline{K}$	$\underline{k}$	$\underline{\Delta/\sigma}$	$\underline{\beta(i)}$	$\underline{E}_U$	$\underline{E}_R$	$\underline{E}_A$
240	16	2	$\beta(0)$	10.0	59.0	100.0
240	16	2	$\beta(8)$	9.9	58.5	94.1
240	16	4	$\beta(0)$	10.0	59.0	100.0
240	16	4	$\beta(8)$	9.9	58.5	94.1
240	24	2	$\beta(0)$	13.2	66.8	100.0
240	24	2	$\beta(12)$	12.9	65.6	91.4
240	24	4	$\beta(0)$	13.2	66.8	100.0
240	24	4	$\beta(12)$	12.9	65.6	91.4
240	32	2	$\beta(0)$	16.3	74.1	100.0
240	32	2	$\beta(16)$	15.8	71.9	89.0
240	32	4	$\beta(0)$	16.3	74.1	100.0
240	32	4	$\beta(16)$	15.8	71.9	89.0

Table 4. Performance Results for GS(3, 0.05907, 0.55223) Strategy. All Results Are Expressed As Percentages.

In Sections 4 and 5 we compare the "optimal" RB strategy with the "optimal" GS strategy for a number of experimental conditions. We found that the optimal GS strategy is generally better than the optimal RB strategy at low Type I error rates but begins to lose its advantage as one considers screening at higher Type I error rates. For example, at a controlled Type I error rate of 10%, the optimal RB strategy is better than the optimal GS strategy when expected relative testing cost is less than (approximately) 45%, at least over the conditions we examined.

Determination of the optimal GS strategy, however, requires prior knowledge of the number of active factors, the signal-to-noise ratio of the active effects, and the directions of the active effects. The RB screening technique requires no such prior knowledge. We discuss the effects of imprecise prior knowledge on the group screening method in Section 6. The analysis of this section indicates that inaccurate prior knowledge can have undesirable consequences on screening performance in that one cannot control the resulting Type I error rate and expected relative testing cost. This is a major drawback to the use of the GS strategy as a technique for factor screening. More importantly, this apparent lack of "robustness" severely limits the practicality of the GS strategy. It remains to be seen, however, how the RB and GS strategies compare in the framework where the active effects are not necessarily assumed to be of the same absolute magnitude. We consider this more general situation in Part II of this technical report.

8. APPENDIX

Listed below and on the following page are the optimal GS strategies as determined by computer-aided search routine for the experimental conditions described in Section 4. The values of α_1 and α_2 are rounded to five decimal places.

K	k	Δ/σ	E_U	E(R)	GS.0	GS.[k/2]
60	4	2	10%	12	GS(30, 0.00747, 1.00000)	GS(30, 0.01811, 1.00000)
				26	GS(7, 0.01097, 0.59142)	GS(7, 0.01447, 0.57278)
				38	GS(5, 0.02248, 0.38328)	GS(4, 0.01200, 0.51311)
				52	GS(3, 0.26978, 0.27305)	GS(2, 0.16950, 0.44325)
60	4	2	20%	12	GS(30, 0.00747, 1.00000)	GS(30, 0.01811, 1.00000)
				26	GS(7, 0.01097, 1.00000)	GS(7, 0.01447, 1.00000)
				38	GS(5, 0.02248, 0.76656)	GS(4, 0.01200, 1.00000)
				52	GS(3, 0.26978, 0.54610)	GS(2, 0.16950, 0.88649)
60	16	2	10%	12	GS(30, 0.00188, 1.00000)	GS(30, 0.00927, 1.00000)
				26	GS(12, 0.00078, 0.45991)	GS(12, 0.00432, 0.42828)
				38	GS(7, 0.00325, 0.29673)	GS(12, 0.01265, 0.23161)
				52	GS(7, 0.00930, 0.17649)	GS(12, 0.04630, 0.14983)
60	16	2	20%	12	GS(30, 0.00188, 1.00000)	GS(30, 0.00927, 1.00000)
				26	GS(12, 0.00078, 0.91982)	GS(12, 0.00432, 0.85657)
				38	GS(7, 0.00325, 0.59346)	GS(12, 0.01265, 0.46323)
				52	GS(7, 0.00930, 0.35298)	GS(12, 0.04630, 0.29848)
60	4	4	10%	12	GS(30, 0.00375, 1.00000)	GS(30, 0.00947, 1.00000)
				26	GS(7, 0.00289, 0.59421)	GS(7, 0.00387, 0.57601)
				38	GS(5, 0.01153, 0.38570)	GS(4, 0.01021, 0.51404)
				52	GS(2, 0.16698, 0.44752)	GS(2, 0.16911, 0.44358)
60	4	4	20%	12	GS(30, 0.00375, 1.00000)	GS(30, 0.00947, 1.00000)
				26	GS(7, 0.00289, 1.00000)	GS(7, 0.00387, 1.00000)
				38	GS(5, 0.01153, 0.77139)	GS(4, 0.01021, 1.00000)
				52	GS(2, 0.16698, 0.89504)	GS(2, 0.16911, 0.88716)
60	16	4	10%	12	GS(30, 0.00094, 1.00000)	GS(30, 0.00470, 1.00000)
				26	GS(12, 0.00019, 0.45997)	GS(12, 0.00109, 0.42858)
				38	GS(7, 0.00082, 0.29697)	GS(12, 0.00324, 0.23166)
				52	GS(7, 0.00235, 0.17660)	GS(12, 0.01238, 0.14986)
60	16	4	20%	12	GS(30, 0.00094, 1.00000)	GS(30, 0.00470, 1.00000)
				26	GS(12, 0.00019, 0.91993)	GS(12, 0.00109, 0.85716)
				38	GS(7, 0.00082, 0.59394)	GS(12, 0.00324, 0.46332)
				52	GS(7, 0.00235, 0.35319)	GS(12, 0.01238, 0.29971)

K	k	Δ/σ	E_U	E(R)	GS.0	GS.[k/2]
240	16	2	10%	46	GS(22, 0.00001, 0.98824)	GS(20, 0.00027, 0.94384)
				100	GS(6, 0.00083, 0.53956)	GS(6, 0.00113, 0.52542)
				144	GS(3, 0.05907, 0.55223)	GS(3, 0.06633, 0.53966)
				198	GS(2, 0.18920, 0.41071)	GS(2, 0.19125, 0.40751)
240	16	2	20%	46	GS(22, 0.00001, 1.00000)	GS(20, 0.00027, 1.00000)
				100	GS(6, 0.00083, 1.00000)	GS(6, 0.00113, 1.00000)
				144	GS(3, 0.05907, 1.00000)	GS(3, 0.06633, 1.00000)
				198	GS(2, 0.18920, 0.82143)	GS(2, 0.19125, 0.81502)
240	64	2	10%	46	GS(27, 0.00005, 0.82463)	GS(27, 0.00052, 0.79563)
				100	GS(16, 0.00000, 0.35317)	GS(27, 0.00228, 0.28641)
				144	GS(9, 0.00001, 0.25538)	GS(27, 0.00599, 0.18749)
				198	GS(9, 0.00004, 0.16109)	GS(40, 0.10492, 0.12894)
240	64	2	20%	46	GS(27, 0.00005, 1.00000)	GS(27, 0.00052, 1.00000)
				100	GS(16, 0.00000, 0.70634)	GS(27, 0.00228, 0.57282)
				144	GS(9, 0.00001, 0.51075)	GS(27, 0.00599, 0.37497)
				198	GS(9, 0.00004, 0.32217)	GS(40, 0.10492, 0.25788)
240	16	4	10%	46	GS(22, 0.00000, 0.98796)	GS(20, 0.00003, 0.94505)
				100	GS(6, 0.00011, 0.53985)	GS(6, 0.00014, 0.52574)
				144	GS(3, 0.05907, 0.55223)	GS(3, 0.06633, 0.53966)
				198	GS(2, 0.18920, 0.41071)	GS(2, 0,19125, 0.40751)
240	16	4	20%	46	GS(22, 0.00000, 1.00000)	GS(20, 0.00003, 1.00000)
				100	GS(6, 0.00011, 1.00000)	GS(6, 0.00014, 1.00000)
				144	GS(3, 0.05907, 1.00000)	GS(3, 0.06633, 1.00000)
				198	GS(2, 0.18920, 0.82143)	GS(2, 0.19125, 0.81502)
240	64	4	10%	46	GS(27, 0.00001, 0.82351)	GS(27, 0.00013, 0.79592)
				100	GS(16, 0.00000, 0.35314)	GS(27, 0.00057, 0.28646)
				144	GS(9, 0.00000, 0.25543)	GS(27, 0.00151, 0.18748)
				198	GS(9, 0.00000, 0.16111)	GS(40, 0.05425, 0.12894)
240	64	4	20%	46	GS(27, 0.00001, 1.00000)	GS(27, 0.00013, 1.00000)
				100	GS(16, 0.00000, 0.70629)	GS(27, 0.00057, 0.57292)
				144	GS(9, 0.00000, 0.51086)	GS(27, 0.00015, 0.37496)
				198	GS(9, 0.00000, 0.32222)	GS(40, 0.05425, 0.25789)

9. REFERENCES

1. Kleijnen, J.P.C. (1975). "Screening Designs for Poly-Factor Experimentation," Technometrics, Vol. 18, pp. 487-493.

2. Mauro, C.A. (1983). "On the Performance of Two-Stage Group Screening Experiments," Desmatics, Inc. Technical Report No. 113-9.

3. Mauro, C.A. (1983). "A Note On The Analysis of Random Balance Designs," Desmatics, Inc. Technical Report No. 113-11.

4. Mauro, C.A., and Burns, K.C. (1983). "Two-Stage Group Screening Search Routine: Program Description and User's Guide," Desmatics, Inc. Technical Report No. 113-10.

5. Mauro, C.A., and Smith, D.E. (1982). "The Performance of Two-Stage Group Screening in Factor Screening Experiments," Technometrics, Vol. 24, pp. 325-330.

6. Mauro, C.A., and Smith, D.E. (1984). "Factor Screening in Simulation: Evaluation of Two Strategies Based on Random Balance Sampling," Management Science, Vol. 30, No. 2 (to appear).

7. Montgomery, D.C. (1979). "Methods for Factor Screening in Computer Simulation Experiments," Georgia Institute of Technology Technical Report.

8. Plackett, R.L. and Burman, J.P. (1946). "The Design of Optimum Multifactor Experiments," Biometrika, Vol. 33, pp. 305-325.

9. Smith, D.E., and Mauro, C.A. (1982). "Factor Screening In Computer Simulation," Simulation, Vol. 38, pp. 49-54.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 113-12	2. GOVT ACCESSION NO. AD A136 487	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Comparison of Random Balance and Two-Stage Group Screening Designs - Part I		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Carl A. Mauro		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0650
9. PERFORMING ORGANIZATION NAME AND ADDRESS Desmatics, Inc. P.O. Box 618 State College, PA 16801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-467
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE December 1983
		13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this report is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Factor Screening
Computer Simulation
Group Screening
Random Balance

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report is the first of two reports which compare the efficacy and relative merits of a two-stage group screening strategy versus a random balance screening strategy. These techniques for factor screening have potential application in the study of large-scale simulation models.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

LMED

8